Available online at www.sciencedirect.com
ScienceDirect
Journal of Multivariate Analysis 99 (2008) 735–750
Journal of Multivariate Analysis
www.elsevier.com/locate/jmva

Generalized Bayes minimax estimators of location vectors for spherically symmetric distributions

Dominique Fourdrinier[a], William E. Strawderman[b,*]

[a] *UMR CNRS 6085, Université de Rouen, LITIS, BP 12, 76801 Saint-Étienne-du-Rouvray, France*
[b] *Hill Center, Department of Statistics, Rutgers University, 110 Frelinghuysen Rd., Piscataway, NJ 08854, USA*

Received 18 August 2006
Available online 5 April 2007

Abstract

Let $X \sim f(\|x-\theta\|^2)$ and let $\delta_\pi(X)$ be the generalized Bayes estimator of θ with respect to a spherically symmetric prior, $\pi(\|\theta\|^2)$, for loss $\|\delta-\theta\|^2$. We show that if $\pi(t)$ is superharmonic, non-increasing, and has a non-decreasing Laplacian, then the generalized Bayes estimator is minimax and dominates the usual minimax estimator $\delta_0(X)=X$ under certain conditions on $f()$. The class of priors includes priors of the form $\left(\frac{1}{A+\|\theta\|^2}\right)^k$ for $k \leqslant \frac{p}{2}-1$ and hence includes the fundamental harmonic prior $\frac{1}{\|\theta\|^{p-2}}$. The class of sampling distributions includes certain variance mixtures of normals and other functions $f(t)$ of the form $e^{-\alpha t^\beta}$ and $e^{-\alpha t+\beta\phi(t)}$ which are not mixtures of normals. The proofs do not rely on boundness or monotonicity of the function $r(t)$ in the representation of the Bayes estimator as $\delta_\pi(X)=\left(1-\frac{ar(t)}{t}\right)X$.

AMS 1991 subject classification: primary 62C10;62C20

Keywords: Minimax estimators; Bayes estimators; Quadratic loss; Spherically symmetric distributions; Location parameter; Superharmonic priors

1. Introduction

This paper is devoted to the study of minimaxity of generalized Bayes estimators, under quadratic loss, of the location parameter of a spherically symmetric distribution. More specifically let X be a random vector in $\mathbb{R}^p$ with a density spherically symmetric about an

* Corresponding author. Fax: +1 732 445 3428.
E-mail address: straw@stat.rutgers.edu (W.E. Strawderman).

0047-259X/$ - see front matter © 2007 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmva.2007.03.007

unknown vector $\theta \in \mathbb{R}^p$, that is,

$$X \sim f\left(\|x-\theta\|^2\right). \tag{1}$$

Consider a generalized prior

$$\pi(\theta) \tag{2}$$

and quadratic loss given by

$$\|\delta-\theta\|^2. \tag{3}$$

We assume throughout that $E_0(\|X\|^2) < \infty$.

The generalized Bayes estimator is the posterior mean and is given by

$$\delta_\pi(X) = X + \frac{1}{m(X)} \int_{\mathbb{R}^p} (\theta - X) f(\|X-\theta\|^2)\pi(\theta)\, d\theta, \tag{4}$$

where $m(X)$ is the marginal

$$m(X) = \int_{\mathbb{R}^p} f(\|X-\theta\|^2)\pi(\theta)\, d\theta. \tag{5}$$

It is well known since Stein [12] that, in the Gaussian case ($f(t) \propto e^{-t/2\sigma^2}$ with σ^2 known), the superharmonicity of $\sqrt{m(X)}$ is a sufficient condition for minimaxity of $\delta_\pi(X)$. This superharmonicity is implied by that of $m(X)$ and in turn by that of $\pi(\theta)$.

In the non-normal case, minimaxity has been studied by many authors (for example, see [13,1–3]). Most of the literature consider Baranchik-type estimators of the form $X - \frac{ar(\|X\|^2)}{\|X\|^2}X$ and derive minimaxity for a in a finite interval through boundedness and monotonicity properties of $r(t)$. Only a few study minimaxity of Bayes estimators. These papers include Strawderman [13], Maruyama [10], Fourdrinier et al. [9]. These authors also establish minimaxity through the Baranchik representation and study only sampling distributions f which are variance mixtures of normal distributions.

In this paper, we establish minimaxity of generalized Bayes estimators for broad classes of spherically symmetric distributions which are not restricted to variance mixtures of normals. Minimaxity results are obtained for unimodal spherically symmetric superharmonic priors $\pi(\|\theta\|^2)$ under the additional condition that the Laplacian of $\pi(\|\theta\|^2)$ is non-decreasing. This class includes priors of the form $(\|\theta\|^2 + A)^{-k}$ with $0 \leqslant k \leqslant p/2 - 1$ and $A \geqslant 0$, and hence contains the fundamental harmonic prior $\|\theta\|^{2-p}$.

Our conditions on the sampling density are that $\frac{f'(t)}{f(t)}$ is non-decreasing and $\frac{F(t)}{f(t)} \geqslant c > 0$, where $F(t) = \frac{1}{2}\int_t^\infty f(u)\, du$ in addition to

$$\int_0^\infty f(t) t^{p/2}\, dt \leqslant 4c \int_0^\infty -f'(t) t^{p/2}\, dt < \infty.$$

This class includes variance mixtures of normal distributions satisfying mild conditions on the variance mixing distribution and many other spherical distributions.

We do not rely on the Baranchik approach mentioned above and our results are more in the spirit of Stein [12].

In Section 2, we develop the model and give preliminary risk calculations. Section 3 is devoted to the main result. Section 4 gives examples of priors and of sampling distributions for which our result holds. In Section 5, we give some concluding remarks and the appendix contains the proofs of some technical lemmas.

2. Bayes estimators and risk functions

Suppose the density, prior and loss are given, respectively, by (1), (2) and (3). Define, for any $t \geqslant 0$, the function $F(t)$ by

$$F(t) = \frac{1}{2} \int_t^\infty f(u)\, du \tag{6}$$

and, for any $x \in \mathbb{R}^p$,

$$M(x) = \int_{\mathbb{R}^p} F(\|x-\theta\|^2)\pi(\theta)\, d\theta. \tag{7}$$

Then the generalized Bayes estimator $\delta_\pi(X)$ defined by (4) and (5) can be written as

$$\delta_\pi(X) = X + \frac{\nabla M(X)}{m(X)}, \tag{8}$$

where ∇ denotes the gradient.

Since X is minimax, it will follow that $\delta_\pi(X)$ itself is minimax provided the difference in risk

$$\Delta_\theta(\delta_\pi) = E_\theta\left[\|\delta_\pi(X)-\theta\|^2 - \|X-\theta\|^2\right] \tag{9}$$

is non-positive. To this end consider the risk difference between a general estimator $\delta(X) = X + g(X)$ and X

$$\Delta_\theta(\delta) = E_\theta\left[\|X+g(X)-\theta\|^2 - \|X-\theta\|^2\right] = E_\theta\left[2(X-\theta)\cdot g(X) + \|g(X)\|^2\right]. \tag{10}$$

The following lemma is useful (see also [1]). Our proof is based on Stokes theorem.

Lemma 2.1. *Let g be a weakly differentiable function from $\mathbb{R}^p$ into $\mathbb{R}^p$. Then*

$$E_\theta\left[(X-\theta)\cdot g(X)\right] = E_\theta\left[\frac{F(\|X-\theta\|^2)}{f(\|X-\theta\|^2)} \operatorname{div} g(X)\right] \tag{11}$$

provided these expectations exist.

Proof. We have

$$\begin{aligned} E_\theta\left[(X-\theta)\cdot g(X)\right] &= \int_{\mathbb{R}^p} (x-\theta)\cdot g(x) f(\|x-\theta\|^2)\, dx \\ &= \int_0^\infty \int_{S_{r,\theta}} (x-\theta)\cdot g(x)\, d\sigma_{r,\theta}(x) f(r^2)\, dr, \end{aligned}$$

where $\sigma_{r,\theta}$ is the uniform measure on the sphere $S_{r,\theta}$ of radius r centered at θ. Introducing the unit normal exterior vector $\eta(x) = \frac{x-\theta}{\|x-\theta\|}$ and using the Divergence Theorem (or Gauss–Green or Stokes theorem, see, e.g. [7, Theorem 4.5.6]) in the form $\int_{S_{r,\theta}} \eta(x) \cdot g(x)\, d\sigma_{r,\theta}(x) = \int_{B_{r,\theta}} \operatorname{div} g(x)\, dx$, we obtain

$$\begin{aligned} E_\theta\left[(X-\theta)\cdot g(X)\right] &= \int_0^\infty \int_{S_{r,\theta}} \frac{x-\theta}{\|x-\theta\|} \cdot g(x)\, d\sigma_{r,\theta}(x) r f(r^2)\, dr \\ &= \int_0^\infty \int_{B_{r,\theta}} \operatorname{div} g(x)\, dx r f(r^2)\, dr \\ &= \int_{\mathbb{R}^p} \int_{\|x-\theta\|}^\infty r f(r^2)\, dr \operatorname{div} g(x)\, dx \end{aligned}$$

by Fubini's theorem. By a change of variable the above equals

$$\int_{\mathbb{R}^p} \frac{1}{2} \int_{\|x-\theta\|^2}^\infty f(u)\, du \operatorname{div} g(x)\, dx = \int_{\mathbb{R}^p} F\left(\|x-\theta\|^2\right) \operatorname{div} g(x)\, dx$$

which is equivalent to (11). □

The next lemma follows immediately from Lemma 2.1.

Lemma 2.2. *The risk difference in* (10) *equals*

$$\Delta_\theta(\delta) = E_\theta\left[2\frac{F(\|X-\theta\|^2)}{f(\|X-\theta\|^2)} \operatorname{div} g(X) + \|g(X)\|^2\right]. \tag{12}$$

A sufficient condition for minimaxity of estimators $\delta(X)$ with finite risk is given by the following lemma when $p \geqslant 3$.

Lemma 2.3. *Assume that* $\frac{F(t)}{f(t)} \geqslant c > 0$ *for all* $t \geqslant 0$ *and that* $E_\theta\left[\|g(X)\|^2\right] < \infty$ *for all* $\theta \in \mathbb{R}^p$. *Then the estimator* $\delta(X)$ *is minimax provided*

$$2c \operatorname{div} g(x) + \|g(x)\|^2 \leqslant 0 \tag{13}$$

for all $x \in \mathbb{R}^p$.

The proof follows immediately from Lemma 2.2. Note that $E_\theta[\|g(X)\|^2] < \infty$ is a finiteness risk condition since X has finite risk.

Applying Lemma 2.3 to the Bayes estimator $\delta_\pi(X)$ in (8) gives the following result.

Lemma 2.4. *Assume that* $\frac{F(t)}{f(t)} \geqslant c > 0$ *for all* $t \geqslant 0$ *and that* $E_\theta \left\| \frac{\nabla M(X)}{m(X)} \right\|^2 < \infty$. *Then the Bayes estimator* $\delta_\pi(X)$ *in* (8) *is minimax provided*

$$2c\frac{\Delta M(x)}{m(x)} - 2c\frac{\nabla M(x)\cdot \nabla m(x)}{m^2(x)} + \frac{\|\nabla M(x)\|^2}{m^2(x)} \leqslant 0 \tag{14}$$

for all $x \in \mathbb{R}^p$.

Proof. The lemma follows from the expression of $\operatorname{div}\left(\frac{\nabla M(x)}{m(x)}\right)$. □

Now we assume that the prior in (2) is spherically symmetric, that is, with an abuse of notation, is of the form

$$\pi(\|\theta\|^2). \tag{15}$$

Note that it follows from the spherical symmetry of π that, for any $x \in \mathbb{R}^p$, $m(x)$ and $M(x)$ are functions of $t = \|x\|^2$. With an additional abuse of notation, we denote

$$m(x) = m(t) \quad \text{and} \quad M(x) = M(t)$$

so that

$$\nabla m(x) = 2xm'(t) \quad \text{and} \quad \nabla M(x) = 2xM'(t). \tag{16}$$

The next lemma shows that $M'(t) \leqslant 0$ for unimodal π.

Lemma 2.5. *Assume that $\pi'(t) \leqslant 0$ for any $t \geqslant 0$. Then it follows that $M'(t) \leqslant 0$ for any $t \geqslant 0$.*

The proof of Lemma 2.5 is given in the appendix.

3. The main result

In preparation for our main result, we give several lemmas whose proofs are deferred to the appendix.

Lemma 3.1. *Let π be a prior of the form* (15). *For any $x \in \mathbb{R}^p$,*

$$x \cdot \nabla m(x) = -2 \int_0^\infty H(u, \|x\|^2) u^{p/2} f'(u)\, du$$

and

$$x \cdot \nabla M(x) = \int_0^\infty H(u, \|x\|^2) u^{p/2} f(u)\, du,$$

where, for $u \in \mathbb{R}_+$,

$$H(u, \|x\|^2) = \lambda(B) \int_{B_{\sqrt{u},x}} x \cdot \theta \pi'(\|\theta\|^2)\, dV_{\sqrt{u},x}(\theta) \tag{17}$$

and $V_{\sqrt{u},x}$ is the uniform distribution on the ball $B_{\sqrt{u},x}$ of radius $\sqrt{u}$ centered at x and $\lambda(B)$ is the volume of the unit ball.

Lemma 3.2. *For any $x \in \mathbb{R}^p$, the function $H(u, \|x\|^2)$ in* (17) *is non-decreasing in u provided that $\Delta\pi(\|\theta\|^2)$ is non-decreasing in $\|\theta\|^2$.*

In the next lemma, we adapt the result used in the proof of Theorem 2.1 given in Cellier et al. [5].

Lemma 3.3. *Let $h(\|\theta - x\|^2)$ be a unimodal density and let $\psi(\theta)$ be a symmetric function. Then*

$$\int_{\mathbb{R}^p} x \cdot \theta \psi(\theta) h(\|\theta - x\|^2)\, d\theta \geqslant 0 \tag{18}$$

as soon as ψ is non-negative.

Our main result is given in the following theorem.

Theorem 3.1. *Suppose X has a spherically symmetric distribution in* $\mathbb{R}^p$ *with density* $f(\|x-\theta\|^2)$. *Suppose* $\theta \in \mathbb{R}^p$ *has a superharmonic prior* $\pi(\|\theta\|^2)$ *such that* $\pi(\|\theta\|^2)$ *is non-increasing and* $\Delta\pi(\|\theta\|^2)$ *is non-decreasing in* $\|\theta\|^2$. *Then the Bayes estimator* δ_π *is minimax under quadratic loss* (3) *provided that* $\frac{f'(t)}{f(t)}$ *is non-decreasing,* $\frac{F(t)}{f(t)} \geqslant c > 0$ *for all* $t \geqslant 0$ *and*

$$\int_0^\infty f(t)t^{p/2}\,dt \leqslant 4c \int_0^\infty -f'(t)t^{p/2}\,dt < \infty. \tag{19}$$

Proof. By superharmonicity of $\pi(\|\theta\|^2)$, we have $\Delta M(x) \leqslant 0$ for all $x \in \mathbb{R}^p$ so that, by Lemma 2.4, it suffices to prove that

$$-2c\nabla M(x)\cdot\nabla m(x) + \|\nabla M(x)\|^2 \leqslant 0 \tag{20}$$

for all $x \in \mathbb{R}^p$. Since m and M are spherically symmetric, by (16), (20) reduces to

$$-2cM'(t)m'(t) + \left(M'(t)\right)^2 \leqslant 0, \tag{21}$$

where $t = \|x\|^2$. Since $M'(t) \leqslant 0$ by Lemma 2.5, Inequality (20) reduces to

$$-2cm'(t) + M'(t) \geqslant 0$$

or, by (16), to

$$-2cx\cdot\nabla m(x) + x\cdot\nabla M(x) \geqslant 0. \tag{22}$$

Note that, by Lemma 3.1, (22) may be expressed as

$$4c\int_0^\infty H(u,t)u^{p/2}f'(u)\,du + \int_0^\infty H(u,t)u^{p/2}f(u)\,du \geqslant 0$$

or

$$4cE\left[H(u,t)\,\frac{f'(u)}{f(u)}\right] + E\left[H(u,t)\right] \geqslant 0, \tag{23}$$

where E denotes the expectation with respect to the density proportional to $u^{p/2}f(u)$. Since, by assumption, $\Delta\pi(\|\theta\|^2)$ is non-decreasing in $\|\theta\|^2$, $H(u,t)$ is non-decreasing in u by Lemma 3.2. Furthermore $\frac{f'(u)}{f(u)}$ is non-decreasing by assumption so that Inequality (23) is satisfied as soon as

$$4cE\left[H(u,t)\right]E\left[\frac{f'(u)}{f(u)}\right] + E\left[H(u,t)\right] \geqslant 0. \tag{24}$$

Finally, as $\pi'(\|\theta\|^2) \leqslant 0$ by assumption, Lemma 3.3 guarantees that $H(u,t) \leqslant 0$ (note that $V_{B,x}$ has a unimodal density) and hence (24) reduces to

$$4cE\left[\frac{f'(u)}{f(u)}\right] + 1 \leqslant 0$$

which is equivalent to (19). $\square$

Remark 3.1. It is worth noting that, since $\frac{f'(t)}{f(t)}$ is non-decreasing, it follows that $\frac{F(t)}{f(t)}$ is non-decreasing as well and hence the lower bound c equals $\frac{F(0)}{f(0)}$. Indeed we can write

$$\frac{f(t)}{F(t)} = 2\frac{\int_t^\infty \frac{-f'(u)}{f(u)} f(u)\,du}{\int_t^\infty f(u)\,du} = 2\,E_t\left[\frac{-f'(u)}{f(u)}\right],$$

where E_t denotes the expectation with respect to the density proportional to $f(u)\mathbf{1}_{]t,\infty[}(u)$ which has increasing monotone likelihood ratio. Hence the monotonicity follows.

4. Examples

4.1. Examples of priors

Conditions on the prior distributions $\pi(\|\theta\|^2)$ of Theorem 3.1 can be expressed as

(a) $\pi'(\|\theta\|^2) \leqslant 0$,
(b) $\Delta\pi(\|\theta\|^2) = 2\left(p\pi'(\|\theta\|^2) + 2\|\theta\|^2\pi''(\|\theta\|^2)\right) \leqslant 0$,
(c) $\frac{d}{d\|\theta\|^2}\,\Delta\pi(\|\theta\|^2) = 2\left((p+2)\pi''(\|\theta\|^2) + 2\|\theta\|^2\pi'''(\|\theta\|^2)\right) \geqslant 0$.

In this section, we give four classes of priors for which (a)–(c) are satisfied. It is interesting to note that Condition (b) implies Condition (c) for each of these examples. Hence the additional Condition (c) on the monotonicity of the Laplacian of $\pi(\|\theta\|^2)$ appears to be a weak condition.

Example (P1). Priors related to the fundamental harmonic prior.

Consider $\pi(t) = \left(\frac{1}{A+t}\right)^k$ with $A \geqslant 0$ and $k \geqslant 0$. It is easy to see that, for $m \in \mathbb{N}$,

$$\pi^{(m)}(t) = \frac{(-1)^m k(k+1)\cdots(k+m-1)}{(A+t)^{k+m}}.$$

Condition (a) is obvious. As for Condition (b), we have

$$\frac{-pk}{(A+t)^{k+1}} + \frac{2k(k+1)t}{(A+t)^{k+2}} \leqslant \frac{-pk + 2k(k+1)}{(A+t)^{k+1}} \leqslant 0$$

for $k \leqslant \frac{p}{2} - 1$. Similarly Condition (c) is also satisfied for $k \leqslant \frac{p}{2} - 1$ since

$$\frac{(p+2)k(k+1)}{(A+t)^{k+2}} - \frac{2k(k+1)(k+2)t}{(A+t)^{k+3}} \geqslant \frac{(p+2)k(k+1) - 2k(k+1)(k+2)}{(A+t)^{k+2}} \geqslant 0.$$

Hence (a)–(c) are satisfied for $0 \leqslant k \leqslant \frac{p}{2} - 1$ and $A \geqslant 0$. Note that the case where $A = 0$ and $k = \frac{p}{2} - 1$ corresponds to the fundamental harmonic prior. We note that the class of priors also arises as a scale mixture of normals (P3). Because of its fundamental importance we have elected to present it separately.

Example (P2). Mixtures of priors satisfying (a)–(c).

Let $(\pi_\alpha)_{\alpha\in A}$ be a family of priors such that Conditions (a)–(c) are satisfied for any $\alpha \in A$. It is clear that the corresponding monotonicity conditions are preserved under any mixture of the form $\int_A \pi_\alpha(\|\theta\|^2)\,dH(\alpha)$ where H is a probability on A.

Consider, for instance, example (P1) with $k=1$, $p \geqslant 4$, $A=\alpha$ and the gamma density $\alpha \mapsto \frac{\beta^{1-v}}{\Gamma(1-v)} \alpha^{-v} e^{-\beta\alpha}$ with $\beta > 0$ and $0 < v < 1$. The integral

$$\int_0^\infty \frac{1}{\alpha + t} \frac{\beta^{1-v}}{\Gamma(1-v)} \alpha^{-v} e^{-\beta\alpha} \, d\alpha$$

is the Stieltjes transform of that density and equals, according to formula (17) of Chapter XIV, Section 2 of Erdélyi [6],

$$t^{-v} e^{\beta t} \Gamma(v, \beta t),$$

where

$$\Gamma(v, \beta t) = \int_{\beta t}^\infty e^{-x} x^{v-1} \, dx. \tag{25}$$

Thus the prior $\|\theta\|^{-2-v} e^{\beta \|\theta\|^2} \Gamma(v, \beta \|\theta\|^2)$ satisfies (a)–(c).

Example (P3). Variance mixtures of normals.

For simplicity, consider the mixture of normals with respect to the inverse of the variance

$$\pi\left(\|\theta\|^2\right) = \int_0^\infty \left(\frac{u}{2\pi}\right)^{p/2} \exp\left(\frac{-u\|\theta\|^2}{2}\right) h(u) \, du.$$

It is clear that, for any $m \in \mathbb{N}$,

$$(2\pi)^{p/2} \pi^{(m)}(t) = \left(-\frac{1}{2}\right)^m \int_0^\infty u^{p/2+m} \exp\left(-\frac{ut}{2}\right) h(u) \, du$$

so that Condition (a) is obvious and Condition (b) reduces to

$$\int_0^\infty u^{p/2+2} \exp\left(-\frac{ut}{2}\right) h(u) \, du \leqslant \frac{p}{t} \int_0^\infty u^{p/2+1} \exp\left(-\frac{ut}{2}\right) h(u) \, du \tag{26}$$

for $t > 0$.

Assuming that $\lim_{u \to 0,\infty} u^{p/2+2} \exp\left(-\frac{ut}{2}\right) h(u) = 0$, an integration by parts expresses the left-hand side of (26) as

$$\frac{2}{t} \int_0^\infty \left(\frac{p}{2} + 2\right) u^{p/2+1} \exp\left(-\frac{ut}{2}\right) h(u) \, du + \frac{2}{t} \int_0^\infty u^{p/2+2} \exp\left(-\frac{ut}{2}\right) h'(u) \, du.$$

Then (26) is equivalent to

$$\int_0^\infty \left[4 + 2\frac{uh'(u)}{h(u)}\right] u^{p/2+1} \exp\left(-\frac{ut}{2}\right) h(u) \, du \leqslant 0$$

which is satisfied as soon as

$$\frac{uh'(u)}{h(u)} \leqslant -2. \tag{27}$$

By a similar integration by parts, it can be shown that the latter condition guarantees Condition (c) as well. The equality case in (27) corresponds to $h(u) \propto \frac{1}{u^2}$ and yields the fundamental harmonic prior $\frac{1}{\|\theta\|^{p-2}}$.

The variance mixtures of normals have been considered in Strawderman [13], Maruyama [10] and Fourdrinier et al. [9]. The latter two papers impose a monotone likelihood ratio property on h which is equivalent to monotonicity of $\frac{uh'(u)}{h(u)}$. Here we do not require this monotonicity but we assume the bound in (27). Also note that, in the latter two papers, the priors can be proper while, here, the assumption of superharmonicity does not allow $\pi(\|\theta\|^2)$ to be proper. This lack of propriety is compensated by the fact that our results apply to a more general class of sampling distributions.

Example (P4). A constructive approach.

Other priors can be obtained through the choice of a non-positive, non-increasing and continuously differentiable function φ such that

$$\int_0^\infty -\varphi(u)u^{p/2-1}\,du < \infty. \tag{28}$$

To any such function φ corresponds a prior $\pi(t)$ defined for all $t > 0$ satisfying

$$\pi'(t) = t^{-p/2}\left(\int_0^t \varphi(u)u^{p/2-1}\,du + k\right), \tag{29}$$

where k is a constant which can be chosen, in view of (28), such that $\pi'(t)\leqslant 0$. Then the prior is expressed as

$$\pi(t) = \int_t^\infty -v^{-p/2}\left(\int_0^v \varphi(u)u^{p/2-1}\,du + k\right)dv. \tag{30}$$

Note that (28) and (29) and the fact that $\pi'(t)\leqslant 0$ imply that $0\leqslant\pi(t) < \infty$ for $t > 0$. It is straightforward to check that $\pi(t)$ satisfies Conditions (a)–(c).

As an example, let $\varphi(u) = -\alpha e^{-\beta u}u^\delta$ with $\alpha\geqslant 0$, $\beta > 0$ and $-p/2 < \delta\leqslant 0$. According to (29), it is easy to check that

$$\pi'(t) = t^{-p/2}\left(\frac{-\alpha}{\beta^{\delta+p/2}}\,\gamma(\delta + p/2, \beta t) + k\right),$$

where

$$\gamma(v, u) = \int_0^u e^{-v}v^{v-1}\,dv \tag{31}$$

is the lower incomplete gamma function and $k\leqslant 0$. Now (30) gives rise to the prior

$$\pi(t) = \int_t^\infty v^{-p/2}\left(\frac{\alpha}{\beta^{\delta+p/2}}\gamma(\delta + p/2, \beta v) - k\right)dv.$$

It can be shown through Fubini's theorem that

$$\pi(t) = \frac{\alpha}{\beta^{\delta+p/2}}\left[t^{-p/2+1}\gamma(\delta + p/2, t\beta) + \frac{1}{\beta^{-p/2+1}}\,\Gamma(\delta + 1, t\beta)\right] + \frac{k}{-p/2+1}t^{-p/2+1},$$

where $\gamma(v, u)$ and $\Gamma(v, u)$ are, respectively, given by (31) and (25).

Note that, for $\alpha = 0$ and $k < 0$, π is the fundamental harmonic function.

4.2. Examples of sampling distributions

Recall that conditions of Theorem 3.1 on the sampling distribution of $f(\|x-\theta\|^2)$ require that

(d) $\frac{f'(t)}{f(t)}$ is non-decreasing,

(e) $\frac{F(t)}{f(t)} \geqslant c > 0$,

(f) $\int_0^\infty f(t) t^{p/2}\, dt \leqslant 4c \int_0^\infty -f'(t) t^{p/2}\, dt < \infty$.

We give three classes of sampling distributions satisfying (d)–(f).

Example (S1). Variance mixtures of normals.

Assume that

$$f(t) = K \int_0^\infty v^{-p/2} \exp\left(-\frac{t}{2v}\right) h(v)\, dv, \tag{32}$$

where h is a mixing density and $K = (2\pi)^{-p/2}$. Note that $\frac{f'(t)}{f(t)}$ is always non-decreasing and $\frac{F(t)}{f(t)} \geqslant \frac{F(0)}{f(0)}$ so that (d) and (e) are satisfied provided that $\frac{F(0)}{f(0)} > 0$. Indeed we have

$$\begin{aligned}\frac{f'(t)}{f(t)} &= \frac{-\dfrac{1}{2}\int_0^\infty v^{-p/2-1} \exp\left(-\dfrac{t}{2v}\right) h(v)\, dv}{\int_0^\infty v^{-p/2} \exp\left(-\dfrac{t}{2v}\right) h(v)\, dv} \\ &= -\frac{1}{2} E_t[V^{-1}],\end{aligned}$$

where E_t denotes the expectation with respect to the density $f_t(v) \propto v^{-p/2} \exp\left(-\frac{t}{2v}\right) h(v)$ which has increasing monotone likelihood ratio in t. Since $-V^{-1}$ is increasing, the monotonicity of $\frac{f'(t)}{f(t)}$ follows.

Similarly, it is easy to see that

$$\frac{F(t)}{f(t)} = E_t[V]$$

and hence is also increasing (see also Remark 3.1). Therefore

$$\frac{F(t)}{f(t)} \geqslant \frac{F(0)}{f(0)} = \frac{\int_0^\infty v^{-p/2+1} h(v)\, dv}{\int_0^\infty v^{-p/2} h(v)\, dv} = \frac{E[V^{-p/2+1}]}{E[V^{-p/2}]} = c. \tag{33}$$

The finiteness of $E[V^{-p/2}]$ guarantees $c > 0$ in (e).

We are now ready to give general conditions for Bayes estimators to be minimax.

Corollary 4.1. *Assume that a prior π satisfies Conditions* (a)–(c) *of Section* 4.1. *Assume also that $f(t)$ is given by* (32) *such that $E[V^{-p/2}] < \infty$. Then the corresponding generalized Bayes estimator is minimax provided that*

$$\frac{E[V]E[V^{-p/2}]}{E[V^{-p/2+1}]} \leqslant 2. \tag{34}$$

Proof. It remains to only verify Condition (f). First, according to (32) and applying Fubini's theorem,

$$\begin{aligned}\int_0^\infty f(t)t^{p/2}\,dt &= K\int_0^\infty\int_0^\infty \exp\left(\frac{-t}{2v}\right)t^{p/2}\,dt\,v^{-p/2}h(v)\,dv\\ &= K\Gamma(p/2+1)2^{p/2+1}\int_0^\infty vh(v)\,dv.\end{aligned}$$

Similarly

$$\begin{aligned}\int_0^\infty -f'(t)t^{p/2}\,dt &= \frac{K}{2}\int_0^\infty\int_0^\infty \exp\left(\frac{-t}{2v}\right)t^{p/2}\,dt\,v^{-p/2-1}h(v)\,dv\\ &= K\Gamma(p/2+1)2^{p/2}\end{aligned}$$

since h is a density. Then Condition (f) is equivalent to $E[V]\leqslant 2c$ which, by (33), reduces to (34). □

Perhaps the most common mixing distribution encountered in practice is the inverse gamma. Assume that $\frac{1}{V} \sim$ Gamma (α, β) with $\alpha > 0$ and $\beta > 0$. The only condition required on the mixing distribution is (34). Here, for $m > -\alpha$,

$$E[V^{-m}] = \frac{\Gamma(\alpha+m)}{\Gamma(\alpha)\beta^m}.$$

So (34) is expressed as

$$\frac{\beta}{\alpha-1}\leqslant 2\beta\frac{\Gamma(\alpha+p/2-1)}{\Gamma(\alpha+p/2)} = 2\frac{\beta}{\alpha+p/2-1},$$

that is,

$$p/2+1\leqslant\alpha.$$

Note first that the scale parameter β plays no role for any mixing distribution. In particular, if the distribution of X is a multivariate t with k degrees of freedom corresponding to $\alpha = \beta = k/2$, Condition (34) reduces to $k\geqslant p+2$.

Example (S2). Densities proportional to $e^{-\alpha t^\beta}$.

Assume that

$$f(t) = Ke^{-\alpha t^\beta},$$

where $\alpha > 0$, $0 < \beta\leqslant 1$ and K is the normalizing constant.

Note that

$$\frac{f'(t)}{f(t)} = -\alpha\beta t^{\beta-1}$$

and is non-decreasing for the range of values of β.

Now let $G(\delta) = \int_0^\infty u^\delta e^{-\alpha u^\beta}\,du$. Through the change of variable $v = \alpha u^\beta$ it can be checked that, for $\delta > -1$,

$$G(\delta) = \left(\frac{1}{\alpha}\right)^{(1+\delta)/\beta}\frac{1}{\beta}\Gamma\left(\frac{\delta+1}{\beta}\right). \tag{35}$$

Then, by Remark 3.1,

$$c = \frac{F(0)}{f(0)} = \frac{1}{2} G(0) = \frac{1}{2} \left(\frac{1}{\alpha} \right)^{1/\beta} \frac{1}{\beta} \Gamma \left(\frac{1}{\beta} \right) > 0. \tag{36}$$

As for Condition (f), in terms of the function G we have

$$\int_0^\infty f(t) t^{p/2} \, dt = KG \left(\frac{p}{2} \right)$$

and

$$\int_0^\infty -f'(t) t^{p/2} \, dt = \alpha \beta K G \left(\beta - 1 + \frac{p}{2} \right).$$

Then Condition (f) is expressed as

$$G \left(\frac{p}{2} \right) \leqslant 4 c \alpha G \left(\beta - 1 + \frac{p}{2} \right),$$

that is, according to (35) and (36),

$$\begin{aligned} &\left(\frac{1}{\alpha} \right)^{(1+p/2)/\beta} \frac{1}{\beta} \Gamma \left(\frac{p/2+1}{\beta} \right) \\ &\quad \leqslant 2 \left(\frac{1}{\alpha} \right)^{1/\beta} \frac{1}{\beta} \Gamma \left(\frac{1}{\beta} \right) \times \alpha \beta \times \left(\frac{1}{\alpha} \right)^{(\beta+p/2)/\beta} \frac{1}{\beta} \Gamma \left(\frac{\beta + p/2}{\beta} \right) \end{aligned}$$

which, after simplification, reduces to

$$\frac{\Gamma \left(\dfrac{p/2+1}{\beta} \right)}{\Gamma \left(\dfrac{1}{\beta} \right) \Gamma \left(\dfrac{\beta + p/2}{\beta} \right)} \leqslant 2. \tag{37}$$

It is easy to see that (37) is satisfied for β in a neighborhood of the form $]1 - \epsilon, 1]$ with $\epsilon > 0$ and not satisfied for $\beta = \frac{1}{2}$.

Example (S3). Densities proportional to $e^{-\alpha t + \beta \varphi(t)}$.

Assume that

$$f(t) = K \exp \left(-\alpha t + \beta \varphi(t) \right),$$

where $\alpha > 0$, $\beta > 0$, $\varphi(0) < \infty$, $\varphi(t) \geqslant 0$, $\varphi'(t) \leqslant 0$, $\varphi'' \geqslant 0$, and

$$\int_0^\infty e^{-\alpha t} t^{p/2} |\varphi'(t)| \, dt < \infty.$$

Note that

$$\frac{f'(t)}{f(t)} = -\alpha + \beta \varphi'(t)$$

and is non-decreasing since the function φ is convex.

By Remark 3.1,

$$0 < c = \frac{F(0)}{f(0)} = \frac{1}{2} \frac{\int_0^\infty \exp(-\alpha u + \beta\varphi(u))\, du}{\exp(\beta\varphi(0))} \leqslant \frac{1}{2\alpha}$$

by assumption on φ.

Now

$$\int_0^\infty f(t) t^{p/2}\, dt \leqslant K \exp(\beta\varphi(0)) \frac{\Gamma(p/2+1)}{\alpha^{p/2+1}} < \infty$$

and then

$$\int_0^\infty -f'(t) t^{p/2}\, dt \leqslant K \exp(\beta\varphi(0)) \left\{ \frac{\Gamma(p/2+1)}{\alpha^{p/2}} + \beta \int_0^\infty \exp(-\alpha t) t^{p/2} |\varphi'(t)|\, dt \right\} < \infty.$$

Note that Condition (f) reduces to

$$\int_0^\infty f(t) t^{p/2}\, dt \leqslant 4c\alpha \int_0^\infty f(t) t^{p/2}\, dt + 4c\beta \int_0^\infty -\varphi'(t) f(t) t^{p/2}\, dt$$

and a simple sufficient condition for minimaxity of the corresponding Bayes estimator is $4c\alpha \geqslant 1$.

The following guarantees minimaxity for any $\alpha > 0$ and for any function φ satisfying the above conditions and for $\beta \leqslant \frac{\log 2}{\varphi(0) - \varphi(1/\alpha)}$. Indeed, according to the expression of c, we have

$$c = \frac{1}{2\alpha} \frac{1}{(\beta\varphi(0))} E_\alpha \left[\exp(\beta\varphi(u)) \right],$$

where E_α denotes the expectation with respect to the density $\alpha e^{-\alpha u}$. As the function $\exp(\beta\varphi(u))$ is convex, we have

$$c \geqslant \frac{1}{2\alpha} \frac{1}{\exp(\beta\varphi(0))} \exp\left(\beta\varphi\left(\frac{1}{\alpha}\right)\right)$$

by Jensen inequality. Hence $4c\alpha \geqslant 1$ as soon as

$$2 \exp\left(\beta \left(\varphi\left(\frac{1}{\alpha}\right) - \varphi(0) \right)\right) \geqslant 1$$

which is equivalent to $\beta \leqslant \frac{\log 2}{\varphi(0) - \varphi(1/\alpha)}$.

5. Concluding remarks

For a broad class of spherically symmetric distributions, we have demonstrated minimaxity of generalized Bayes estimators corresponding to a superharmonic prior under mild additional conditions on the prior. This subclass contains variance mixtures of normal distributions under finite moment conditions and large subclasses which are not mixtures of normals.

The approach in this paper is different from most of the literature for spherically symmetric distributions in that we do not rely on monotonicity and boundedness properties of the shrinkage function in the usual Baranchik representation. The use of the superharmonicity of the prior is more in the spirit of Stein [12].

Recently Maruyama [10] and Fourdrinier et al. [9] gave classes of priors including some proper priors for which the resulting Bayes estimators are minimax in the case where both the sampling distributions and priors are mixtures of normals. Note that, due to the superharmonicity property, our priors cannot be proper (see [8]). However, neither our priors nor our densities need to be mixture of normals. Furthermore, in the case of mixtures of normal sampling distributions, our mixing distributions are not required to have monotone likelihood ratio as in the above papers.

A conjecture of Brown [4] suggests that generalized Bayes estimators for priors $\pi(\theta) \sim \|\theta\|^a$ with $a \leqslant 2 - p$ are admissible. Hence it appears that generalized Bayes estimators for priors in our class such that $\pi(\theta) \sim \|\theta\|^{2-p}$ are admissible minimax. In our setting, a recent result of Maruyama and Takemura [11] supports this conjecture under additional assumptions.

Appendix A.

Proof of Lemma 2.5. It follows from (16) that the sign of $M'(t)$ is the same as the sign of $x \cdot \nabla M(x)$. As noted in the proof of Theorem 3.1 (after (24)), the function $H(u, t)$ defined in (17) is non-positive. Hence, by Lemma 3.1, $x \cdot \nabla M(x) \leqslant 0$ and $M'(t) \leqslant 0$. $\square$

Proof of Lemma 3.1. According to (5), we have

$$\begin{aligned} x \cdot \nabla m(x) &= 2 \int_{\mathbb{R}^p} x \cdot (x - \theta) f'(\|x - \theta\|^2) \pi(\|\theta\|^2)\, d\theta \\ &= 2 \int_0^\infty \int_{S_{R,x}} x \cdot (x - \theta) \pi(\|\theta\|^2)\, d\sigma_{R,x}(\theta) f'(R^2)\, dR, \end{aligned}$$

where $\sigma_{R,x}$ is the uniform measure on the sphere $S_{R,x}$ of radius R centered at x.

Through $\frac{\theta - x}{\|\theta - x\|}$, the unit normal exterior vector at $\theta \in S_{R,x}$, we have

$$\begin{aligned} x \cdot \nabla m(x) &= 2 \int_0^\infty \int_{S_{R,x}} -\pi(\|\theta\|^2) x \cdot \frac{\theta - x}{\|\theta - x\|}\, d\sigma_{R,x}(\theta) R f'(R^2)\, dR \\ &= -2 \int_0^\infty \int_{B_{R,x}} \mathrm{div}_\theta \left(\pi(\|\theta\|^2) x\right) d\theta R f'(R^2)\, dR \end{aligned}$$

by Stokes theorem and hence

$$\begin{aligned} x \cdot \nabla m(x) &= -2 \int_0^\infty \int_{B_{R,x}} x \cdot \nabla_\theta \left(\pi(\|\theta\|^2)\right) d\theta R f'(R^2)\, dR \\ &= -4 \int_0^\infty \int_{B_{R,x}} x \cdot \theta \pi'(\|\theta\|^2)\, d\theta R f'(R^2)\, dR \\ &= -4\lambda(B) \int_0^\infty \int_{B_{R,x}} x \cdot \theta \pi'(\|\theta\|^2)\, dV_{R,x}(\theta) R^{p+1} f'(R^2)\, dR \end{aligned}$$

according to the definition of $V_{R,x}$. Then

$$\begin{aligned} x \cdot \nabla m(x) &= -4 \int_0^\infty H(R^2, \|x\|^2) R^{p+1} f'(R^2)\, dR \\ &= -2 \int_0^\infty H(u, \|x\|^2) u^{p/2} f'(u)\, du \end{aligned}$$

through the change of variable $u = R^2$.

This is the first result. The second result follows in the same way referring to (7). $\square$

The following lemma is needed for the proof of Lemma 3.2. It is known but, for completeness, we give a proof using Stokes theorem.

Lemma A.1. *If $g(\theta)$ is a twice continuously differentiable function on $\mathbb{R}^p$ then, for any $x \in \mathbb{R}^p$,*

$$\frac{d}{dR}\int_{S_{R,x}} g(\theta)\, dU_{R,x}(\theta) = \frac{R}{p}\int_{B_{R,x}} \Delta g(\theta)\, dV_{R,x}(\theta).$$

Proof. Through a change of variable we have

$$\begin{aligned}\frac{d}{dR}\int_{S_{R,x}} g(\theta)\, dU_{R,x}(\theta) &= \frac{d}{dR}\int_{S} g(R\eta + x)\, dU(\eta)\\ &= \int_S \frac{\partial}{\partial R} g(R\eta + x)\, dU(\eta)\\ &= \frac{1}{\sigma(S)}\int_S \nabla g(R\eta + x)\cdot \eta\, d\sigma(\eta).\end{aligned}$$

Then, by Stokes theorem,

$$\begin{aligned}\frac{d}{dR}\int_{S_{R,x}} g(\theta)\, dU_{R,x}(\theta) &= \frac{1}{\sigma(S)}\int_B \operatorname{div}\left(\nabla g(R\eta + x)\right)\, d\eta\\ &= \frac{R}{\sigma(S)}\int_B \Delta g(R\eta + x)\, d\eta\\ &= \frac{R}{\sigma(S)}\int_{B_{R,x}} \frac{\Delta g(\theta)}{R^p}\, d\theta\\ &= \frac{R}{p}\int_{B_{R,x}} \Delta g(\theta)\, dV_{R,x}(\theta). \qquad \square\end{aligned}$$

Proof of Lemma 3.2. The result will follow from the monotonicity in R of

$$\int_{S_{R,x}} x\cdot\theta\pi'(\|\theta\|^2)\, dU_{R,x}(\theta) \tag{38}$$

since

$$\begin{aligned}\int_{B_{R,x}} x\cdot\theta\pi'(\|\theta\|^2)\, dV_{R,x}(\theta) &= \frac{p}{R^p}\int_0^R \tau^{p-1}\int_{S_{\tau,x}} x\cdot\theta\pi'(\|\theta\|^2)\, dU_{\tau,x}(\theta)\, d\tau\\ &= p\int_0^1 u^{p-1}\int_{S_{Ru,x}} x\cdot\theta\pi'(\|\theta\|^2)\, dU_{Ru,x}(\theta)\, du.\end{aligned}$$

Now deriving (38) with respect to R gives through Lemma A.1

$$\frac{d}{dR}\int_{S_{R,x}} x\cdot\theta\pi'(\|\theta\|^2)\, dU_{R,x}(\theta) = \frac{R}{p}\int_{B_{R,x}} \Delta\left(x\cdot\theta\pi'(\|\theta\|^2)\right)\, dV_{R,x}(\theta).$$

Since (see the expressions in Conditions (b) and (c) in Section 4.1)

$$\Delta\left(x\cdot\theta\pi'(\|\theta\|^2)\right) = x\cdot\theta\varphi'(\|\theta\|^2),$$

where

$$\Delta\pi(\|\theta\|^2) = \varphi(\|\theta\|^2),$$

we have

$$\frac{d}{dR}\int_{S_{R,x}} x \cdot \theta\pi'(\|\theta\|^2)\, dU_{R,x}(\theta) = \frac{R}{p}\int_{B_{R,x}} x \cdot \theta\varphi'(\|\theta\|^2)\, dV_{R,x}(\theta).$$

By assumption on the monotonicity of the Laplacian of $\pi(\|\theta\|^2)$, we have $\varphi'(\|\theta\|^2) \geqslant 0$ which, by Lemma 3.3, implies that the last integral is non-negative since $V_{R,x}$ is unimodal. $\square$

References

[1] J.O. Berger, Minimax estimation of location vectors for a wide class of densities, Ann. Statist. 3 (1975) 1318–1328.
[2] A. Brandwein, W.E. Strawderman, Minimax estimators of location parameters for spherically symmetric unimodal distributions under quadratic loss, Ann. Statist. 6 (1978) 377–416.
[3] A. Brandwein, W.E. Strawderman, Generalizations of James–Stein estimators under spherical symmetry, Ann. Statist. 19 (1991) 1639–1650.
[4] L.D. Brown, A heuristic method for determining admissibility of estimators with applications, Ann. Statist. 7 (1979) 960–994.
[5] D. Cellier, D. Fourdrinier, W.E. Strawderman, Shrinkage positive rule estimators for spherically symmetric distributions, J. Multivariate Anal. 53 (1995) 194–209.
[6] A. Erdélyi, M.F. Oberhettinger, F.G. Tricomi, in: A. Erdélyi (Ed.), Tables of Integral Transforms, McGraw-Hill, New York, 1954.
[7] H. Federer, Geometric Measure Theory, Springer, Berlin, 1969.
[8] D. Fourdrinier, O. Kortbi, W.E. Strawderman, Bayes minimax estimators of the mean of a scale mixture of multivariate normal distributions, J. Multivariate Anal. (2007), to appear.
[9] D. Fourdrinier, W.E. Strawderman, M.T. Wells, On the construction of Bayes minimax estimators, Ann. Statist. 26 (1998) 660–671.
[10] Y. Maruyama, Admissible minimax estimators of a mean vector of scale mixtures of multivariate normal distribution, J. Multivariate Anal. 21 (2003) 69–78.
[11] Y. Maruyama, A. Takemura, Admissibility and minimaxity of generalized Bayes estimators for spherically symmetric family, Technical Report METR 2005-13, Department of Mathematical Informatics, University of Tokyo, 2006.
[12] C. Stein, Estimation of the mean of multivariate normal distribution, Ann. Math. Statist. 9 (1981) 1135–1151.
[13] W.E. Strawderman, Minimax estimation of location parameters for certain spherically symmetric distribution, J. Multivariate Anal. 4 (1974) 255–264.